A STATISTICAL ANALYSIS TO TEST THE RELIABILITY OF MEASURING ATMOSPHERIC NOISE SUBJECTIVELY BY A SMALL GROUP OF PEOPLE*

B. B. GHOSH

RESEARCH DEPARTMENT, A.I.R., NEW DELHI

(*Received, September* 30, 1959)

ABSTRACT. A statistical analysis of the results of some preliminary listening tests on the variations of judgement of different groups of listeners for the subjective measurement of atmospheric noise has been described. The possible limits of error committed by a comparatively smaller sample size or group of listeners as compared to that of a large group have been indicated and compared with those obtained from theoretical considerations. On the basis of these findings, a practical procedure to meet the statistical requirement in employing a small group of eight observers for the actual measurement of atmospheric noise by a subjective method like that of Thomas, but suitably modified, for broadcasting services has been discussed.

INTRODUCTION

In any experimental investigation where random events are encountered, a statistical approach has to be adopted for obtaining reliable results. Atmospheric noise, as is well known, constitutes one such random phenomenon and statistical methods have to be employed for its measurements. Usually, two methods of measurements are followed :

The objective method in which noise voltage is observed in a meter or recorded graphically or photographed from oscilloscope and the subjective one in which the minimum signal to over-ride the annoying effect of atmospheric noise is assessed by listening to signals of different strengths mixed with noise. In the first method of measurements, statistical analysis of the observation is required to arrive at the median, higher decile and lower decile values of noise. In the second method also, reliability of assessment by individual listener or a group is required to be known from statistical considerations and any error reduced to a minimum.

In an earlier communication Ghosh and Mitra (1958) have discussed both the methods of measurements and presented the measured data on atmospheric noise at Delhi from 1955 onwards. Thomas method of measuring atmospheric noise has been suitably modified to suit the requirements of broadcasting and a

* Communicated by Prof. S. K. Mitra.

statistical correlation established between the objective noise and subjective assessment of minimum signal required to suppress the noise. Adoption of the Thomas method for subjective evaluation of noise for broadcasting purposes naturally calls for statistical considerations of the range of errors incurred in taking individual or group listening as the correct assessment of minimum satisfactory signal. We have employed 8 people whose collective judgement of this signal has been accepted as accurate enough for all practical purposes. Very elaborate statistical tests were conducted before this minimum number of people, required for the purpose without incurring any significant error, was arrived at It is the purpose of this paper to determine the range of errors involved in the choice of such a group of persons and to assess the reliability of such measurements. It will be shown that even such a small group is adequate for giving fairly reliable results on subjective assessment within the limits of experimental error

THEORETICAL CONSIDERATIONS

In the series of tests to determine the size of the group whose collective judgement can be taken as a correct assessment of the minimum satisfactory signal, a number of recorded programmes modulating steady signals which were mixed with atmospheric noise in random proportions, were listened to by a large number

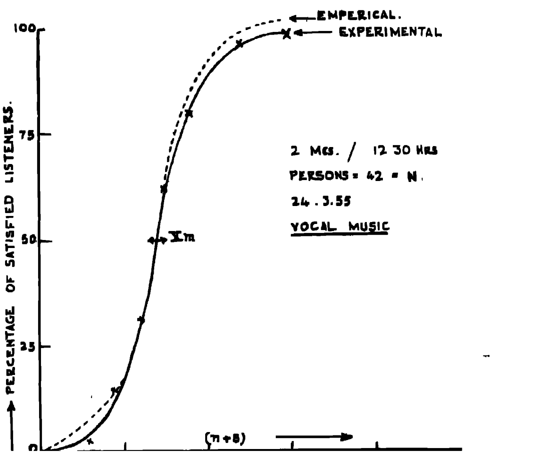

Fig. 1. Shows the variations of the percentage of satisfied listeners, S, for a group of 42 persons against noise (n) puls signal-increases (S) in db. above the initial level.

of listeners in a group. Each recorded piece contained, therefore, a discrete value of signal to noise ratio and each listener was asked to record his opinion as either satisfactory or unsatisfactory from the point of view of annoyance caused by the noise.

Here the limited number of listeners in a group represents a sample of the population or universe consisting of all the listeners to broadcast programmes The characteristic evaluated is the 'estimate' recorded by each individual as either 'satisfactory' or 'unsatisfactory' and is therefore finite and discontinuous. We have sought to determine statistically how, for a given value of S/N ratio, the estimates, random as they are, are distributed. Figure 1 shows the variation of percentage estimate δ with signal increases over a given noise $(n+S)$ for a total number of 42 listeners in one group This is a typical curve for all such distributions where the number of listeners is fairly large and represents a cumulative Gaussian distribution. The curve is of the form represented empirically by

$$\delta = A\,[1 \pm (1 - e^{-\alpha|x|})] \qquad \dots \quad (1)$$

where δ = % of listener-satisfaction

A, α = constants

x = Deviation in db from 50% satisfaction level

It may be noted that near X_m, the median value of $(n+S)$ the curve is very steep indicating that small variation in $(n+S)$ produces large changes in δ.

X_m and σ (standard deviation) for a particular universe have fixed values. But if we consider samples within the universe, the values of X_m and σ may vary from sample to sample.

The distribution of X_m for such samples within the population will be normal. We may, therefore, use standard statistical tables for finding out the range of variation of X_m. It can be shown that if the sample consists of N listeners and the tests are repeated, then there is 95% chance that X_m will lie in the range $\pm \frac{1.96\sigma}{\sqrt{N-1}}$ of its correct value. When N is large, (e.g. 42) repeated listening tests will indicate that X_m for such samples will be very near the median value for the entire universe.

We have calculated this range of variations for X_m for different group sizes using the above criterion. The values of σ used for the computation of X_m have been determined by the usual method of finding the mean and squaring the differences from the mean from curves (like figure 1) actually obtained by group listening experiments. The results are described in the next section.

STATISTICAL ANALYSIS

In the first sample considered, total number of listeners was 42 and the variation of its δ with $(n+S)$ was as shown in figure 1 Its standard deviation σ comes out to be 4.1 and the range of variation of X_m can be found from

$$\overleftrightarrow{X_m} = \pm \frac{2\sigma}{\sqrt{N-1}} \qquad \ldots \quad (2)$$

X_m works out to be $\pm$ 1.28 *db*. If the sample sizo is increased, X_m will be further decreased.

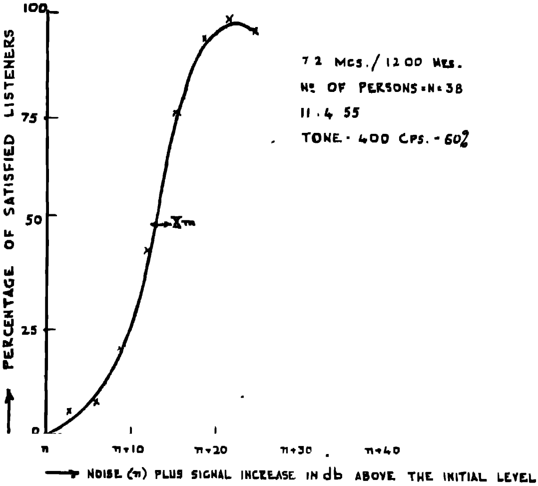

Fig. 2. Shows the variations of δ against $(n+S)$ for a group of 38 persons.

Next we have considered a sample where $N=38$. Its σ comes out from figure 2 to be the same as above and X_m equals $\pm$ 1.3. The experimental results showing the variation of δ are shown in figure 2 for this case and the curve is very similar to figure 1.

Thus, the above two experiments clearly indicate that for large number of listeners of the order of 30 or more, there is not much error involved in the median value of the signal to noise ratio. Our aim, however, is to determine the smallest size of the sample which could be utilised for practical purposes and at the same time the error in estimation by collective listening would not be appreciable. We have, therefore, reduced the size of the sample further and carried out similar analysis.

Figure 3 shows the variation of δ where $N=23$. Its σ is 4.8 and its $\overleftrightarrow{X_m}$ is $\pm$ 2.0 db. Fig. 4 indicates the situation where $N=18$, its σ is 5. 3 and $\overleftrightarrow{X_m}$ is $\pm$ 2.6 dbs. It would be noted from Fig. 4 that the curve has been steeper and slightly unstable. The border line judgement has become more difficult and

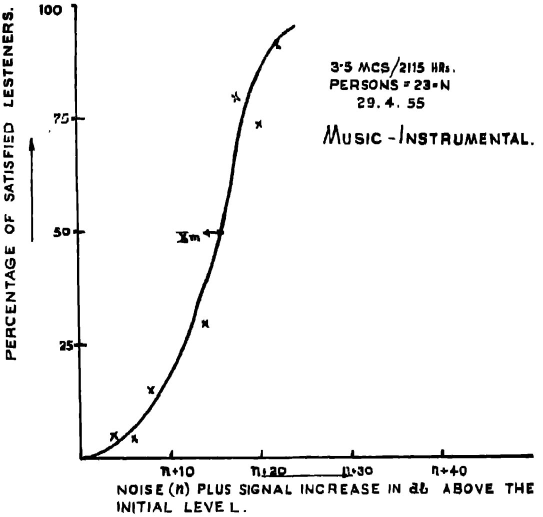

Fig. 3. Shows the variations of δ against (n+S) for a group of 23 persons

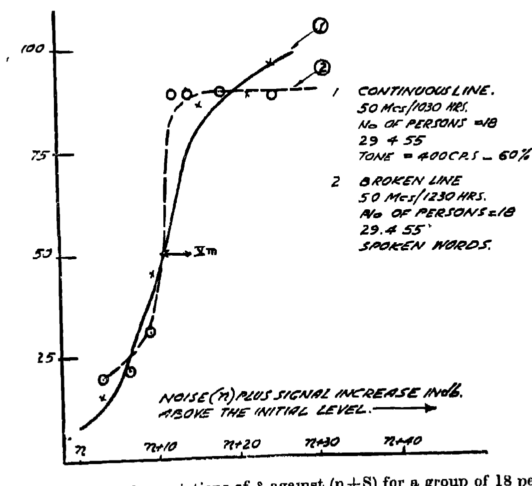

Fig. 4 Shows the variations of δ against (n+S) for a group of 18 persons for two sets of independent group-listening tests.

one or two listeners assessing wrongly have contributed to the change in the shape of the curve at the two ends. This point is illustrated more specifically by the two curves plotted in Fig. 4 corresponding to two independent listening tests. But the significant feature to note is that the median valves for both the listening tests (when $N=18$) have remained the same though the shapes of the two curves are slightly different. To conduct each noise measurement subjectively by 18 people at a time is a difficult task specially when one has to carry out these measurements at several hours during the day and each time on several wave frequencies.

Dinger and Paine (1947) from theoretical arguments have arrived at the conclusion that for a random phenomenon like atmospheric noise any method of measurement should be good enough if the results are repeatable within 30%. Thus if we allow 3 to 4 *db* error in the median values the smallest group size comes out to be about 8, by equation (2), using the above value of $\sigma=5.3$ for $N=18$ for this calculation Actually this value of σ should be slightly higher for lowering $N=18$ to $N=8$

Let us now consider the situation when only 8 persons are employed for the listening tests We have plotted such a curve for 8 from experimental results in figure 5 and the curve is fairly symmetrical about its median and is of the same type as obtained for larger samples.

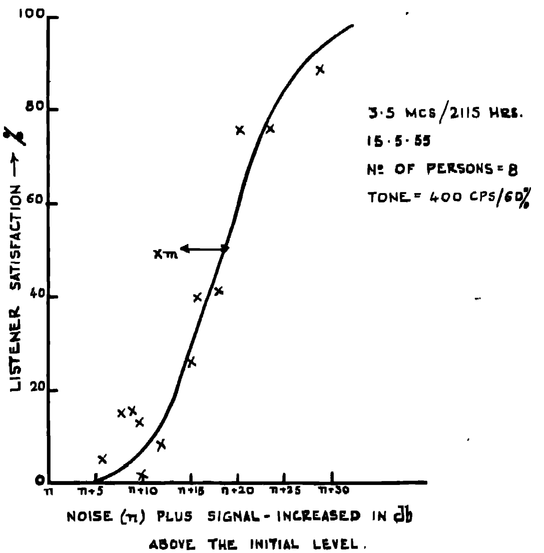

Fig. 5. Shows the variations of δ against (n+S)for a group of 8 persons

σ for this curve is 5.6 about 5% higher than that for $N=18$ and about 35% higher than that for $N=42$ and $\overline{X_m}$ is $+4.3$ *db* Thus even from experimental observations, it is proved that when 8 people are employed for such listening tests, it is expected that there is 95% probability that the median of the satisfactory signal as assessed by this group of listeners will lie within a range of ± 4 to 5 *db* of the correct median value for the universe. In other words, if we can conduct each subjective measurement of atmospheric noise by employing 8 people and take their collective judgement as the measured value of the minimum satisfactory signal, the maximum error we may incur will be about 4 to 5 *db*.

DISCUSSION

We have shown that for subjective measurements of noise by Thomas method when the annoyance to broadcasting programmes is considered, the collective judgement of 8 people would give a maximum error of the order of 4 to 5 *db* in the median value of the minimum satisfactory signal as assessed by them. It should be remembered that no such considerations apply to W/T signals where atmospheric noise is assessed from readability of the signals. Personal errors in estimation are thus minimised and may be within 1 or 2 *db* when experienced persons are employed. But when broadcast programmes consisting of music and talks are used, the assessment of a satisfactory signal in the presence of noise will have a subjective element that may vary from person to person and also on the type of programme used. In fact when 8 people are employed, an incorrect assessment by a single individual will cause a change in the percentage satisfaction by as much as 12.5%. Since these errors in estimation are distributed at random according to a normal law of error, their mean, median or average are the same. Collective judgement of 8 people gives a permissible error in the median (or average) value of the minimum satisfactory signal and therefore, all our subjective measurements have been taken on this basis.

One may ask the question whether an error of 4 to 5 *db* in the average assessment can be termed negligible. It may be remembered that the main object of the noise measurements, as envisaged by us, is in repect of its utilisation for broadcasting purposes. In planning a broadcasting service, we need to know the signal strength laid down at the target area together with the atmospheric noise existing there. In the calculation of field strengths, an accuracy of 4 to 5 *db* is indeed difficult to claim Similarly, if one intends to obtain quantitative estimate of atmospheric noise itself from such subjective assessment, it has already been proved (Ghosh and Mitra—1958) from statistical correlation that a reduction of 40 *db* will give the noise field. Here again the standard deviation is found to be 6 *db*. Thus, an accuracy of 4 to 5 *db* in the assessment of minimum satisfactory signal by a group of 8 persons should be considered adequate for all practical purposes.

We have simplified the problem, without losing further in accuracy in the following way. We know from our experience that at Delhi, noise does not change appreciably from day to day at the same time on the same frequencies, except, of course, during local thunderstorms As we are interested in monthly and seasonal averages, it would be sufficient to take measurement by a single person and change him from day to day so that when the average figure for the month is considered, we have the assessment of a number of individuals. In our receiving centre at Delhi normally 8 to 10 engineers are in duty taking shift at various times of the day. Their individual assessments, when collectively analysed over a month, represent three to four group listening tests of 8 persons in a group. It is reasonable therefore to consider that the range of error is not likely to exceed 4 to 5 *db*, on the other hand, may even be less.

The statistical correlation (Ghosh and Mitra 1958) between subjective assessment and objective measurements of noise taken simultaneously is another proof of the reliability of our method of measurement These two sets of measurements are absolutely independent of each other, even then the figure 40 *db* ($\sigma=6$ *db*) has been found to be the most probable value of the protection needed for a satisfactory signal in the presence of atmosphere noise. Had there been greater variability in the assessment by a group of 8 persons, one would not expect a statistically significant protection ratio with a small standard deviation.

ACKNOWLEDGMENT

The present paper forms a part of programme of propagation research of the Research Department of All India Radio, New Delhi The auother is grateful to his colleagues Mr. P. K. Venkatasubramanian and Mr. S K. Srikanthan for considerable help at the early stages of the investigation

The paper is published by kind permission of the Chief Engineer, All India Radio.

REFERENCES

Dinger, H. E. and Paine, H. G., 1947, "Factor affecting the accuracy of radio noise meters" *P.I.R.E.*, **35**, 75.

Ghosh, B. B. and Mitra, S. N., 1958, "Measurement of Atmospheric Noise", *Jour. I.T.E*, **5**, 2.